

BEST AVAILABLE COPY

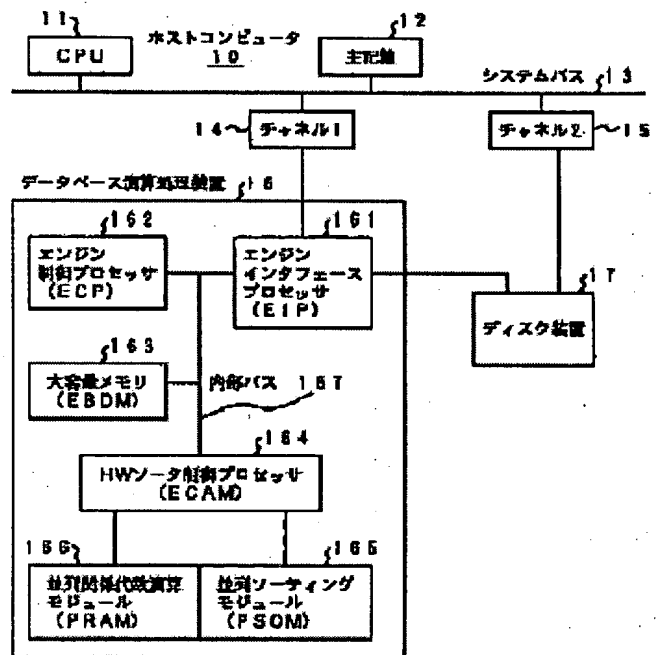
DATA PROCESSOR

Patent number: JP5324580
Publication date: 1993-12-07
Inventor: YAMADA HIROYOSHI, others: 02
Applicant: TOSHIBA CORP
Classification:
 - International: G06F15/16; G06F12/00
 - european:
Application number: JP19920133471 19920526
Priority number(s):

Abstract of JP5324580

PURPOSE: To improve the parallel processing performance by a multi-thread by constituting the processor so that a thread is generated, and also, its processing can be started.

CONSTITUTION: An engine control processor (ECP) 162 divides its arithmetic processing into plural processings which can execute a parallel operation in accordance with an arithmetic request from a host computer 10, and allocates a thread to those processings, by which in a sort processing, a delivery processing of data from a large capacity memory (EBDM) 163 to a hardware sorter control processor (ECAM) 164, and a data reception processing from the hardware sorter control processor (ECAM) 164 are executed in parallel by a multi-thread.



Data supplied from the esp@cenet database - Patent Abstracts of Japan

(19)日本国特許庁(J P)

(12) 公開特許公報(A)

(11)特許出願公開番号

特開平5-324580

(43)公開日 平成5年(1993)12月7日

(51)Int.Cl.*	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 F 15/16	3 7 0 Z	9190-5L		
12/00	5 1 2	8526-5B		

審査請求 未請求 請求項の数2(全 8 頁)

(21)出願番号 特願平4-133471

(22)出願日 平成4年(1992)5月28日

(71)出願人 000003078

株式会社東芝

神奈川県川崎市幸区堀川町72番地

(72)発明者 山田 広佳

東京都青梅市末広町2丁目9番地 株式会
社東芝青梅工場内

(72)発明者 岩崎 孝夫

東京都青梅市末広町2丁目9番地 株式会
社東芝青梅工場内

(72)発明者 島川 和典

東京都青梅市末広町2丁目9番地 株式会
社東芝青梅工場内

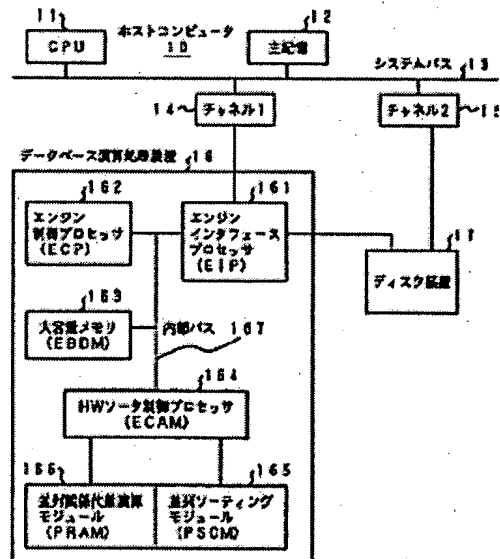
(74)代理人 弁理士 鈴江 武彦

(54)【発明の名称】 データ処理装置

(57)【要約】

【目的】スレッドの生成と共にその処理を開始できるようにし、マルチスレッドによる並行処理性能の向上を図る。

【構成】エンジン制御プロセッサ(ECP)162は、ホストコンピュータ10からの演算要求にしたがってその演算処理を並行動作可能な複数の処理に分割し、それら処理にスレッドを割り当てることにより、ソート処理においては、大容量メモリ(EBDM)163からハードウェアソータ制御プロセッサ(ECAM)164へのデータの引き渡し処理と、ハードウェアソータ制御プロセッサ(ECAM)164からのデータ受信処理とがマルチスレッドによって並行して実行される。



BEST AVAILABLE COPY

【特許請求の範囲】

【請求項 1】 ホスト装置と、演算対象のファイルデータが格納される2次記憶装置と、前記2次記憶装置を直接アクセスするためのバスを有し、前記ホスト装置からの要求に応じて前記演算対象のファイルデータに対して所定の演算処理を実行する演算処理装置とを備えたデータ処理装置において、

前記演算処理装置は、演算対象のファイルデータを前記バスを介して前記2次記憶装置から入力する入力手段と、入力された演算対象のファイルデータに対してソートまたは関係代数演算を実行する演算手段と、前記演算対象のファイルデータおよび前記演算手段による演算結果データが格納される内部メモリと、前記演算結果データを前記2次記憶装置に出力する出力手段と、前記ホスト装置からの演算要求に従って前記演算処理装置による演算処理を制御する制御手段とを具備し、

前記制御手段は、前記ホスト装置から要求された演算処理を互いに並行動作可能な複数の処理に分割し、それら処理毎にスレッドを割り当てることによって前記演算処理をマルチスレッド環境によって並行処理させることを特徴とするデータ処理装置。

【請求項 2】 ホスト装置と、演算対象のファイルデータが格納される2次記憶装置と、前記2次記憶装置を直接アクセスするためのバスを有し、前記ホスト装置からの要求に応じて前記演算対象のファイルデータに対して所定の演算処理を実行する演算処理装置とを備えたデータ処理装置において、

前記演算処理装置は、前記ホスト装置との間の通信、および前記2次記憶装置との間のデータ入出力を実行する第1のプロセッサと、ソートまたは関係代数演算を行う演算回路と、この演算回路による演算を実行制御する第2のプロセッサと、前記第1のプロセッサによって入力される演算対象のファイルデータおよび前記演算回路による演算結果が格納される内部メモリと、前記第1のプロセッサを介して供給される前記ホスト装置からの指示に基づき、前記第1および第2のプロセッサを動作制御する第3のプロセッサとを具備し、

前記第3のプロセッサは、前記ホスト装置から要求された演算処理を互いに独立した複数の処理に分割し、それら処理毎にスレッドを割り当てることによって前記演算処理をマルチスレッド環境によって並行処理させることを特徴とするデータ処理装置。

は、ソート処理や関係データベースの検索系処理等を高速に実行するために専用のデータベース演算処理装置が設けられている。このデータベース演算処理装置はハードウェアと称される演算回路を備えており、この演算回路を用いることによってホスト装置からの演算処理要求に応じたソートや関係代数演算等を高速に実行する。

【0003】このような制御を高速に行うためには、各処理単位を並列に実行することが必要になる。そのため全体を制御するプロセッサはマルチスレッド環境を提供し、各処理単位に応じたスレッドを生成することになる。従来、このスレッドの生成には演算処理の過程が考慮されておらず、このため、互いに別のスレッドのデータ結果を初期情報として受け取るまで実際の処理を開始できないといった不具合があった。この場合、初期情報を受信するまではスレッド管理のオーバーヘッドが増加するだけで、実際の並行処理の効果は得られないことになる。

【0004】

【発明が解決しようとする課題】従来では、初期情報が伝達されるまでスレッド管理の無駄が生じており、これによってCPU資源の浪費、実行効率の悪化を招く欠点があった。この発明はこのような点に鑑みてなされたもので、十分に効率の良い並行処理を実現することができるデータ処理装置を提供することを目的とする。

【0005】

【課題を解決するための手段および作用】この発明は、ホスト装置と、演算対象のファイルデータが格納される2次記憶装置と、前記2次記憶装置を直接アクセスするためのバスを有し、前記ホスト装置からの要求に応じて前記演算対象のファイルデータに対して所定の演算処理を実行する演算処理装置とを備えたデータ処理装置において、前記演算処理装置は、演算対象のファイルデータを前記バスを介して前記2次記憶装置から入力する入力手段と、入力された演算対象のファイルデータに対してソートまたは関係代数演算を実行する演算手段と、前記演算対象のファイルデータおよび前記演算手段による演算結果データが格納される内部メモリと、前記演算結果データを前記2次記憶装置に出力する出力手段と、前記ホスト装置からの演算要求に従って前記演算処理装置による演算処理を制御する制御手段とを具備し、前記制御手段は、前記ホスト装置から要求された演算処理を互いに並行動作可能な複数の処理に分割し、それら処理毎にスレッドを割り当てることによって前記演算処理をマルチスレッド環境によって並行処理させることを第1の特徴とする。

【0006】このデータ処理装置においては、演算処理が互いに並行動作可能な複数の処理に分割され、それら処理毎にスレッドが割り当てられるので、各スレッドは他のスレッドの結果等を待つことなく、生成されると同

【発明の詳細な説明】

【0001】

【産業上の利用分野】この発明は、ホスト装置からの要求に応じて演算対象のファイルデータに対してソートまたは関係代数演算等の演算処理を実行する演算処理装置を備えたデータ処理装置に関する。

【0002】

【従来の技術】一般に、コンピュータシステムにおいて

時に処理を開始することができる。したがって、演算処理装置の並行処理効率を大幅に向上させることが可能となる。

【0007】また、この発明は、データベース演算処理装置に第1乃至第3のプロセッサから構成されるマルチプロセッサ構成を採用し、データベース演算処理装置と2次記憶装置間のデータ入出力を第1のプロセッサに、演算回路を用いた演算処理を第2のプロセッサに、そして、全体の制御を第3のプロセッサにそれぞれ機能分散すると共に、ホスト装置からの演算処理とマルチスレッドによる演算制御の並列処理とを第3のプロセッサによって実行させるように構成したことを第2の特徴とする。

【0008】この構成においては、第1乃至第3のプロセッサによって効率よく負荷分散が図れ、しかも第1のプロセッサによるデータ入出力と、第2のプロセッサによる演算処理との並行動作が可能となり、それら並行動作とそれを実現するスレッドを効率的に生成することを第3のプロセッサが実行制御することにより、データベース演算処理装置の動作性能を十分に向上させることができる。

【0009】

【実施例】以下、図面を参照してこの発明の実施例を説明する。

【0010】図1にはこの発明の一実施例に係わるデータ処理装置の全体のシステム構成が示されている。このデータ処理装置は、ホストコンピュータ10と、データベース演算処理装置（データベースエンジン；DBE）16と、磁気ディスク装置17とによって構成されている。ホストコンピュータ10は、CPU11、主記憶装置12、第1および第2のチャンネル装置14、15によって構成されており、これらCPU11、主記憶装置12、およびチャンネル装置14、15はシステムバス13を介して相互接続されている。

【0011】ホストコンピュータ10とデータベース演算処理装置16はチャンネル装置14によって接続され、またホストコンピュータ10と磁気ディスク装置17はチャンネル装置15によって接続されている。さらに、データベース演算処理装置16と磁気ディスク17は専用のアクセスバス18を介して接続されている。

【0012】CPU11は、ホストコンピュータ10全体の制御を司るものであり、データベース演算処理装置16に対してソート処理や関係代数演算等の各種演算処理の実行を依頼する。主記憶装置12には、データベース演算処理装置16に演算対象ファイルや演算内容を指示するためのコマンドが格納される。

【0013】データベース演算処理装置（DBE）16は、CPU11からのコマンドに基づいて演算対象ファイルのデータに対して演算処理を実行するものであり、演算対象ファイルの入力および演算結果の出力のため

に、バス18を介して磁気ディスク装置17を直接的にアクセスする。

【0014】このデータベース演算処理装置（DBE）16は、エンジンインターフェースプロセッサ（EIP）161、エンジン制御プロセッサ（ECP）162、大容量メモリ（EBDM）163、ハードウェアソータ制御プロセッサ（ECAM）164、並列ソーティングモジュール（PSOM）165、および並列関係代数演算モジュール（PRAM）166によって構成されている。

【0015】エンジンインターフェースプロセッサ（EIP）161、エンジン制御プロセッサ（ECP）162、およびハードウェアソータ制御プロセッサ（ECAM）164の3台のプロセッサは内部バス167によって相互接続されており、大容量メモリ（EBDM）163を共有メモリとする密結合のマルチプロセッサを構成している。大容量メモリ（EBDM）163は、3台の各プロセッサの共通のアドレス空間上に配置されている。また、これら3台の各プロセッサ、つまり、エンジンインターフェースプロセッサ（EIP）161、エンジン制御プロセッサ（ECP）162、およびハードウェアソータ制御プロセッサ（ECAM）164には機能分散がなされている。この場合、それぞれの固有の役割を効率よく実行するために、これらの各プロセッサは、密結合ながらそれぞれに適した独立のモジュールによって動作制御されるように構成されている。

【0016】エンジンインターフェースプロセッサ（EIP）161、エンジン制御プロセッサ（ECP）162、およびハードウェアソータ制御プロセッサ（ECAM）164機能分散は、次のようになされている。

【0017】すなわち、エンジンインターフェースプロセッサ（EIP）161は、ホストコンピュータ10とデータベース演算処理装置16間の通信を行うと共に、磁気ディスク装置17のディスクコントローラにもバス18を介して接続されており、磁気ディスク装置17との間でのデータ入出力を制御する。また、エンジンインターフェースプロセッサ（EIP）161は、磁気ディスク装置17にデータを出力する際、出力ファイルの再構成処理も行う。

【0018】ホストコンピュータ10との通信においては、エンジンインターフェースプロセッサ（EIP）161は、第1のチャンネル装置14を介してCPU11から送られてくるコマンドを受信し、それをエンジン制御プロセッサ（ECP）162に送信する。また、エンジンインターフェースプロセッサ（EIP）161は、エンジン制御プロセッサ（ECP）162から送られてくるコマンド結果としてステータスを受信し、それを第1のチャンネル装置14を介してCPU11に返信する。

【0019】磁気ディスク装置17との間のデータ入出

力においては、エンジンインターフェースプロセッサ（EIP）161は、エンジン制御プロセッサ（ECP）162からの入出力要求を受け付け、大容量メモリ（EBDM）163と磁気ディスク装置17との間でデータ転送を行う。その際、エンジン制御プロセッサ（ECP）162からの要求に従い出力ファイルデータの再構成処理も行う。

【0020】エンジン制御プロセッサ（ECP）162は、エンジンインターフェースプロセッサ（EIP）161、大容量メモリ（EBDM）163、およびハードウェアソータ制御プロセッサ（ECAM）164を内部バス167を介して制御するものであり、CPU11からのコマンドが、ソートや関係代数演算などの演算処理コマンドであった場合には、それに対応する各種コマンド処理プロセスを生成・実行したり、最小の処理単位であるスレッドをその初期情報を与えると共に生成・実行する。

【0021】大容量メモリ（EBDM）163は、磁気ディスク装置17から読み出された演算対象のファイルデータ、CPU11から送られてくる各種演算処理コマンド、ハードウェアソータ制御プロセッサ（ECAM）164、並列ソーティングモジュール（PSOM）165、および並列関係代数演算モジュール（PRAM）166による演算処理結果、さらには、エンジン制御プロセッサ（ECP）162による演算結果の統合結果等を格納する共有メモリである。演算対象ファイルデータは大容量メモリ（EBDM）163内の入力バッファ部に格納され、演算処理結果は大容量メモリ（EBDM）163内の出力バッファ部に格納される。

【0022】ハードウェアソータ制御プロセッサ（ECAM）164は、エンジン制御プロセッサ（ECP）162からの指令に基づいて、並列ソーティングモジュール（PSOM）165および並列関係代数演算モジュール（PRAM）166による演算を制御する。

【0023】この場合、ハードウェアソータ制御プロセッサ（ECAM）164は、大容量メモリ（EBDM）163上のデータを並列ソーティングモジュール（PSOM）165に入力し、並列関係代数演算モジュール（PRAM）166から出力された演算結果を大容量メモリ（EBDM）163上に格納するが、並列ソーティングモジュール（PSOM）165へのデータ入力に際しては、まず、キー切り出し処理を実行する。このキー切り出し処理においては、ハードウェアソータ制御プロセッサ（ECAM）164は、演算対象の各レコードから演算に必要な演算対象キーフィールドのみを切り出し、それにレコード識別番号（大容量メモリ（EBDM）163上におけるレコードの先頭アドレス）を付加して並列ソーティングモジュール（PSOM）165へ送出する。

【0024】並列ソーティングモジュール（PSOM）

165は、大容量メモリ（EBDM）163によって駆動され、ソートを並列に実行する専用のハードウェア回路であり、並列関係代数演算モジュール（PRAM）166に接続されている。この並列ソーティングモジュール（PSOM）165は、バイブライニングソータと称されるものであり、2-ウェイマージを行う複数のソートセルをカスケード接続してなる。

【0025】並列関係代数演算モジュール（PRAM）166は、関係データベースにおけるJOIN（結合）やRESTRICT（制約）といった関係代数演算を並列に実行する専用のハードウェア回路であり、並列ソーティングモジュール（PSOM）165からソートされたデータを入力し、演算結果をハードウェアソータ制御プロセッサ（ECAM）164に出力する。ソート処理だけを実行する場合には、並列関係代数演算モジュール（PRAM）166は、最終段のソートセルとして機能する。

【0026】このように、エンジンインターフェースプロセッサ（EIP）161は、エンジン制御プロセッサ（ECP）162、およびハードウェアソータ制御プロセッサ（ECAM）164にはそれぞれ各種演算処理、例えばソート処理やセレクト処理を実行するための機能が分散されている。この場合、これらプロセッサは、大容量メモリ（EBDM）163上のバッファのやりとりを除けば非同期に動作し、平行して各機能を実行する。すなわち、ソートやセレクト等の処理は、通常、入力処理、演算処理、および出力処理から構成されるが、ハードウェアソータ制御プロセッサ（ECAM）164の制御により演算処理は、エンジンインターフェースプロセッサ（EIP）161による演算対象データの入力処理と並行して実行される。また、ハードウェアソータ制御プロセッサ（ECAM）164からの演算結果をエンジン制御プロセッサ（ECP）162によって統合処理する必要がある場合には、その統合処理とエンジンインターフェースプロセッサ（EIP）161による出力処理も並行して実行される。次に、図2を参照して、演算処理に対するスレッドの割り当て手法を説明する。

【0027】例えば、データベース演算処理装置16がソート処理を行う場合、そのソート処理は、大容量メモリ（EBDM）163上のソート対象データをハードウェアソータ制御プロセッサ（ECAM）164に渡す処理（処理1）と、ハードウェアソータ制御プロセッサ（ECAM）164からソート結果を受け取る処理（処理2）とに分割される。これら処理は、互いに並行処理可能な独立した処理であり、それぞれスレッドとして割り当てられる。さらに、大容量メモリ（EBDM）163に2つの入力バッファを設けて、それらの一方のバッファのデータをハードウェアソータ制御プロセッサ（ECAM）164に渡している時に他方のバッファにデータを用意するといったいわゆるダブルバッファ処理を行

うことにより、処理1は2つの独立した処理に分割される。1つは、第1の入力バッファにデータを用意してそれをハードウェアソータ制御プロセッサ（E C A M）164に渡す処理（処理11）であり、もう一つは、第2の入力バッファにデータを用意してそれをハードウェアソータ制御プロセッサ（E C A M）164に渡す処理（処理12）である。これら処理に対しても、処理1に対応するスレッドから生成される新たなスレッドがそれぞれ割り当てられる。同様に、大容量メモリ（E B D M）163に2つの出力バッファを設けて、それらの一方のバッファのデータをマージ等を行う次の演算処理スレッドに渡している間に他方のバッファにハードウェアソータ制御プロセッサ（E C A M）164から出力されるデータを格納することによって、処理2についても、2つの独立した処理に分割できる。1つは、第1の出力バッファにデータを用意してそれを次の演算処理スレッドに渡す処理（処理21）であり、もう一つは、第2の入力バッファにデータを用意してそれを次の演算処理スレッドに渡す処理（処理22）である。これら処理に対しても、処理2に対応するスレッドから生成される新たなスレッドがそれぞれ割り当てられる。

【0028】このようにして、ソータ処理は並行動作可能な4つの処理に分割され、これらそれぞれにスレッドが割り当てられる。これらスレッドは、それぞれ他のスレッドの結果等を待つことなく独立して処理を進められるので、生成されると同時に処理を開始することができ、次に、図3のフローチャートを参照して、スレッド生成時の処理を説明する。

【0029】スレッドは生成されるとスレッドを管理するためのテーブルに登録され、実行開始待ちキューにつながられる（ステップS21）。これは一般にはエンジン制御プロセッサ（E C P）162内のOSの機能として提供される。従来であれば、生成されたスレッドは処理を開始するための初期情報を獲得するまで待ち状態に入る。しかし、これは機能的にはスレッドの初期設定における一部にほかならず、この実施例ではもともと並行して動作可能な処理にスレッドを割り当てているので、スレッド生成時におけるシステムコールの一部として実現される。よって、スレッドは生成されるとともに処理開始可能な状態になっており（ステップS211、S21

2）、改めて初期情報を受け取る必要はない。

【0030】また、このようにして生成されたスレッドが更に他のスレッドを生成する場合も同様である。新しいスレッドを生成することが必要であれば（ステップS22）、システムコールを用いることによって（ステップS23）、スレッド生成に伴う初期設定を行い（ステップS231）、付随して処理開始のための初期情報を与える（ステップS232）ことになる。こうして、スレッドを生成した後は自スレッドの処理を継続し（ステップS24）、自スレッドの処理を終えるまで（ステップS25）までスレッドは存続する。

【0031】以上のように、この実施例においては、エンジン制御プロセッサ（E C P）162の制御によって、演算処理が互いに並行動作可能な複数の処理に分割され、それら処理毎にスレッドが割り当てられるので、各スレッドは他のスレッドの結果等を待つことなく、生成されると同時に処理を開始することができる。また、並行動作可能な複数の処理それぞれについても、それら各処理が内包している並行動作可能な複数の処理に細分化され、それら処理がスレッドとして生成されるので、並行処理効率を大幅に向上させることが可能となる。

【0032】

【発明の効果】以上のようにこの発明によれば、スレッドを生成と共にその実行が可能になり、効率の良いマルチスレッド環境の実現によってシステム全体として効率のよいデータ処理を実現できる。

【図面の簡単な説明】

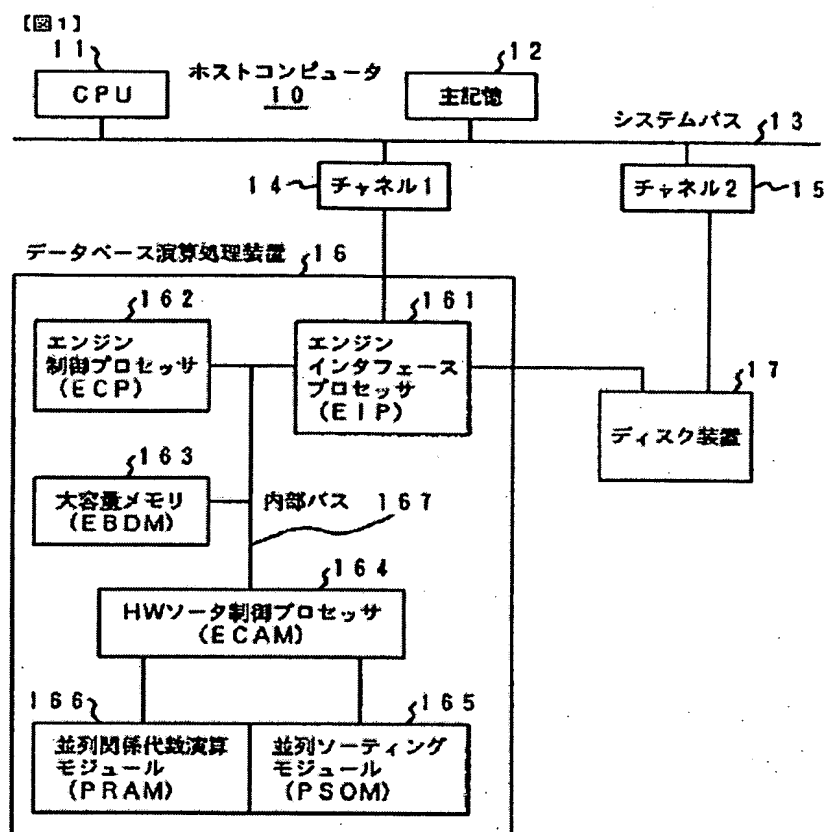
【図1】この発明の一実施例に係る全体のシステム構成を示すブロック図。

【図2】図1のシステムにおける演算処理に対するスレッドの割り当ての一例を示す図。

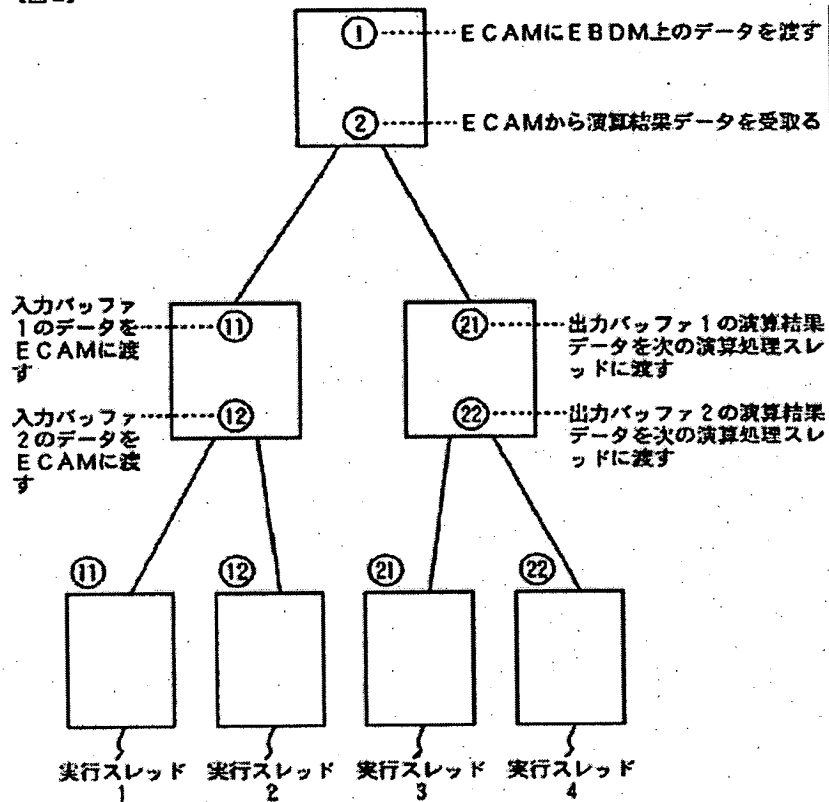
【図3】図1のシステムにおけるスレッド生成動作を説明するためのフローチャート。

【符号の説明】

10…ホストコンピュータ、13…システムバス、14…チャンネル、16…データベース演算処理装置、17…磁気ディスク装置、161、162、164…プロセッサ、163…大容量メモリ、165…並列ソーティングモジュール、166…並列関係代数演算モジュール。



【図2】



【図 3】

